

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

JULIO ANTONIO VAZQUEZ

**APLICAÇÃO DE MODELOS ESTOCÁSTICOS NA ANÁLISE DO
COMPORTAMENTO DE UMA AÇÃO NA BOLSA DE VALORES: UM ESTUDO DE
CASO**

**LONDRINA
2025**

JULIO ANTONIO VAZQUEZ

**APLICAÇÃO DE MODELOS ESTOCÁSTICOS NA ANÁLISE DO
COMPORTAMENTO DE UMA AÇÃO NA BOLSA DE VALORES: UM ESTUDO DE
CASO**

**Application Of Stochastic Models In Analyzing The Behavior Of A Stock
Exchange: A Case Study**

Trabalho de conclusão de curso de graduação
apresentado como requisito para obtenção do título de
Bacharel em Engenharia de Produção da Universidade
Tecnológica Federal do Paraná (UTFPR) do campus
Londrina.

Orientador: Dr. Rogério Tondato.

LONDRINA

2025



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

JULIO ANTONIO VAZQUEZ

**APLICAÇÃO DE MODELOS ESTOCÁSTICOS NA ANÁLISE DO
COMPORTAMENTO DE UMA AÇÃO NA BOLSA DE VALORES: UM ESTUDO DE
CASO**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do título de
Bacharel em Engenharia de Produção da Universidade
Tecnológica Federal do Paraná (UTFPR) do campus
Londrina.

Data de aprovação: 17/fevereiro/2025

Rogério Tondato
Doutor
Universidade Tecnológica Federal do Paraná

Silvana Rodrigues Quintilhano
Doutora
Universidade Tecnológica Federal do Paraná

José Ângelo Ferreira
Doutor
Universidade Tecnológica Federal do Paraná

LONDRINA

2025

Dedico este trabalho à minha família, pelos
momentos de ausência e a todos que me ajudaram
ao longo desta caminhada.

AGRADECIMENTOS

Ao meu orientador Prof. Dr. Rogério Tondato por toda dedicação e confiança em mim depositadas para a elaboração desse trabalho.

A minha família, pelo incentivo e apoio indispensáveis nessa jornada.

Aos meus colegas de faculdade pela amizade, pela troca de experiência e todo o tempo de boa convivência que tivemos.

Também quero agradecer à Universidade Federal Tecnológica do Paraná e a todos os professores do meu curso pela elevada qualidade do ensino oferecido.

RESUMO

O presente trabalho tem por objetivo analisar o comportamento da ação AMZO34 em relação às variáveis AMZN, Dólar e Ibovespa, utilizando ferramentas estatísticas para comprovar a correlação entre elas. Trata-se de uma pesquisa com metodologia de natureza qualitativa e caráter descritivo, que utilizou técnicas estatística e modelos estocásticos, para compreender melhor o mercado financeiro e contribuir para que as pessoas, de modo geral, tenham mais informações sobre a bolsa de valores, podendo tomar decisões mais acertadas para realizar de forma segura um planejamento financeiro. Para desenvolver este trabalho foram coletados e analisados dados sobre as variáveis no período de janeiro a junho de 2024, desconsiderando feriados e finais de semana. A partir do índice de Correlação de Pearson, foi possível identificar que todas as variáveis independentes possuem grau de relação significativo com a variável dependente. Diante disso, foi realizada uma regressão linear múltipla, que proporcionou um modelo altamente significativo, já que o resultado obtido explica de maneira satisfatória a variação dos dados e a força de relação entre as variáveis estudadas.

Palavras-chave: Regressão; Correlação; Séries Temporais; Variáveis.

ABSTRACT

This study aims to analyze the behavior of the AMZO34 stock in relation to the AMZN, Dollar and Ibovespa variables, using statistical tools to prove the correlation between them. This is a study with a qualitative and descriptive methodology, which used statistical techniques and stochastic models, to better understand the financial market and contribute to people, in general, having more information about the stock market, allowing them to make better decisions to safely carry out financial planning. To develop this study, data on the variables were collected and analyzed from January to June 2024, disregarding holidays and weekends. Using Pearson's Correlation index, it was possible to identify that all independent variables have a significant degree of relationship with the dependent variable. Therefore, a multiple linear regression was performed, which provided a highly significant model, since the result obtained satisfactorily explains the variation in the data and the strength of the relationship between the variables studied.

Keywords: Regression; Correlation; Time Series; Variables.

LISTA DE FIGURAS

Figura 1 – Representação gráfica de uma série temporal	17
Figura 2 – Neurônio biológico	18
Figura 3 – Exemplo de árvore de decisão	20
Figura 4 – Gráfico de duas variáveis dependentes	23
Figura 5 – Ajuste linear	23
Figura 6 – Base do método dos mínimos quadrados	24
Figura 7 – AMZO34 x Tempo	29
Figura 8 – AMZN x Tempo	29
Figura 9 – Dólar x Tempo.....	30
Figura 10 – Ibovespa x Tempo	30
Figura 11 – AMZO34 x AMZN.....	31
Figura 12 – AMZO34 x Dólar	32
Figura 13 – AMZO34 x Ibovespa	32
Figura 14 – Resíduos AMZN	35
Figura 15 – Resíduos Dólar	35
Figura 16 – Resíduos Ibovespa.....	35

LISTA DE TABELAS

Tabela 1 – Estatísticas descritivas das variáveis quantitativas	28
Tabela 2 – Coeficiente de Pearson	31
Tabela 3 – Estatísticas de Regressão.....	33
Tabela 4 – ANOVA	33
Tabela 5 – Coeficientes.....	34

SUMÁRIO

1	INTRODUÇÃO	11
1.1	Objetivos	12
1.2	Justificativas.....	12
2	REFERENCIAL TEÓRICO.....	14
2.1	Mercado Financeiro.....	14
2.2	Previsão de Séries Temporais.....	15
2.3	Redes Neurais.....	17
2.4	Árvores de Decisão	19
2.5	Correlação.....	21
2.6	Regressão	22
3	MÉTODOS E TÉCNICAS.....	26
4	RESULTADOS E DISCUSSÕES	28
5	CONSIDERAÇÕES FINAIS	38
	REFERÊNCIAS.....	40

1 INTRODUÇÃO

Atualmente, tem-se discutido sobre a técnica de dados que é cada vez mais importante devido ao mundo que está em constante evolução, onde a informação se tornou um recurso extremamente valioso. Ela desempenha o papel central de guiar decisões estratégicas em diversos setores. A partir da coleta, processamento e interpretação de dados, possibilita a descoberta de tendências, padrões e informações importantes para a otimização do processo, para gerar novas descobertas e conhecimentos, aprimorar a eficiência e promover a inovação. (PRODANOV; FREITAS, 2013)

No presente trabalho, para aplicação das técnicas estatísticas e modelos estocásticos, foram utilizados os valores referentes aos meses de janeiro a junho de 2024, das variáveis AMZO34, AMZN, Dólar e Ibovespa, desconsiderando os finais de semana e feriados, tanto no Brasil quanto nos Estados Unidos, com objetivo de compreender melhor o comportamento da variável AMZO34 e a existência ou não de correlação com as demais variáveis.

A análise dessas variáveis, mesmo que por um curto período de tempo, permite que tanto as pessoas que já atuam no mercado financeiro, quanto as que pretendem ingressar possam avaliar o valor de uma ação, identificar oportunidades e riscos e tomar decisões de investimento.

Conforme Reis (2007), determinados eventos podem ser analisados a partir de séries temporais, que são um conjunto de dados organizados em ordem cronológica, obtidos através de observações ou medições em intervalos regulares ao longo do tempo. Essas séries podem revelar tendências, sazonalidades e padrões que evoluem com o tempo, tornando-se valiosas para previsões e análises em diversas áreas, como finanças, economia, meteorologia e ciências sociais. Ainda segundo o autor, a análise de séries temporais envolve a identificação e modelagem de padrões, permitindo a criação de modelos estatísticos para fazer previsões futuras ou entender melhor o comportamento das variáveis ao longo do tempo.

Outra técnica estatística utilizada para análise é a regressão, em que através dela é possível quantificar a relação entre variáveis dependentes e independentes, permitindo identificar padrões, prever e tomar decisões com base nos dados. Existem diferentes modelos de regressão, incluindo a regressão linear simples e múltipla,

logística, polinomial, sendo que cada uma deve ser utilizada de acordo com as necessidades e objetivos da análise. (VIRGILLITO, 2017)

Os resultados obtidos pela aplicação do modelo de regressão linear múltipla, gerou valores satisfatórios, indicando a boa qualidade do modelo para explicar a variação dos dados e a força de relação entre as variáveis dependente e independentes. Além disso, o valor de F, F significação e os valores-p, mostram que o modelo altamente significativo.

1.1 Objetivos

O objetivo geral deste trabalho é examinar a variação do preço de uma ação nas bolsas de valores Ibovespa e NASDAQ, através de métodos de análise de dados, a fim de verificar a existência de relação estatística.

Como objetivos específicos têm-se:

- Estabelecer uma base teórica sólida abordando os conceitos dessas áreas;
- Compreender a relação sobre os preços das ações e os fatores que os influenciam;
- Coletar dados, envolvendo a obtenção do histórico dos preços das ações em cada bolsa de valores para construir um conjunto de informações;
- Aplicar algoritmos de aprendizagem de máquina para validar a relação;
- Realizar uma análise detalhada dos resultados para determinar se existe ou não alguma relação significativa entre o preço da ação nas diferentes bolsas de valores.

1.2 Justificativas

O preço de uma ação dentro da bolsa de valores sofre muitas variações em um período bem curto de tempo. De acordo com o SERASA (Serviços de Assessoria S.A.) a formação dos preços das ações depende da lei da oferta e da demanda, ou seja, o valor de uma ação depende do número de pessoas que estão interessadas em comprar ou vender tal ação. No entanto, através da análise de dados é possível concluir se existe alguma relação entre o preço de uma ação nas bolsas de valores Ibovespa e NASDAQ. Nesse sentido, essa pesquisa se justifica por contribuir para

que empresas e investidores tenham mais informações sobre o mercado de ações e possam escolher a melhor forma de investir.

De acordo com Prodanov e Freitas (2013), a análise de dados desempenha papel fundamental em uma pesquisa científica, pois através dela é possível transformar informações brutas em conhecimento estruturado, possibilitando interpretações e conclusões confiáveis, sendo relevante para a validação de hipóteses, tomada de decisões, organização e interpretação de informações, precisão e confiabilidade.

A análise de dados pode ser considerado um dos momentos mais críticos da pesquisa científica, pois a forma como os dados são organizados, tratados e interpretados, determina a qualidade dos resultados e conclusões, ou seja, é essencial para que uma pesquisa produza conhecimento válido e útil no meio acadêmico, científico ou corporativo. (PRODANOV; FREITAS, 2013)

2 REFERENCIAL TEÓRICO

2.1 Mercado Financeiro

Nos últimos anos, o número de pessoas que estão atuando no mercado financeiro vem crescendo, principalmente no Brasil. Segundo Guedes (2021) grande parte dos novos investidores da bolsa de valores, cerca de 68%, são jovens e que o público universitário tem dedicado cada vez mais tempo para aprender sobre o mercado financeiro e, por isso, cada vez mais se procura conteúdos relacionados a esse assunto.

Costa (2015, *apud* PAIVA *et al.*, 2020, p.3), cita que “Finanças é a área que trata de assuntos que correspondem ao uso do dinheiro. Sendo assim, a finança pessoal envolve conteúdos relacionados à administração do dinheiro pessoal”.

Atualmente existe uma preocupação muito maior com a área de finança pessoal, já que segundo Ferreira (2008, *apud* PAIVA *et al.*, 2020), o ato de não gastar dinheiro é uma boa forma de conquistar riqueza, no entanto para que haja um crescimento econômico mais acelerado, o ato de investir é um excelente caminho, mas para que isso seja possível é preciso estar sempre atento, pois as oportunidades aparecem com o tempo sendo necessário analisar e decidir qual a melhor opção.

Mesmo com o crescente número de novos ingressantes no mercado financeiro, quando se fala de Brasil, fica evidente um perfil conservador, que investe na poupança para evitar riscos em suas aplicações, sendo que muitos não realizam nenhum tipo de investimento, pela falta de conhecimento na área. (PAIVA, 2020)

Em relação ao endividamento, as principais causas ocorrem pela má administração financeira, compras demasiadas que podem ser pagas por meio do crediário ou cartão de crédito, pagamento rotativo, cheque especial, cheque pré-datado, empréstimo consignado, empréstimo pessoal, financiamentos, ausência de renda mensal devido ao desemprego, entre outros aspectos que interferem na poupança financeira. Portanto, a junção desses fatores pode desencadear a inadimplência. (MEDEIROS, 2020)

Ou seja, a falta de conhecimento sobre administração financeira, além de fazer com que as pessoas deixem de ganhar dinheiro com investimentos, faz com que elas tenham problemas como endividamento e inadimplência. Portanto, a educação financeira é de extrema importância, como cita Medeiros, Silva, Nascimento e Santos (2020, p.214):

A educação financeira abrange diversos impactos, positivos ou negativos, congruente ao perfil das pessoas que adquiriram a autonomia de administrar o próprio capital perante as situações de crises econômicas já ocorridas e lidar com imprevistos financeiros é uma das estratégias para evitar eventuais frustrações e, principalmente, controlar o impulso de gastar sem ter condições de arcar com os prejuízos futuros.

Chambi (2016) assegura que outro motivo que assusta possíveis investidores são as rápidas e constantes mudanças do mercado financeiro, o que o torna de certa forma algo que não é totalmente previsível, e afirma que a volatilidade é uma característica fundamental de todos os mercados financeiros, e para aqueles que operam nele é de vital importância fazer medições e previsões.

Ainda, segundo Chambi (2016), a volatilidade é como a velocidade do mercado é medida, é a rapidez com que os preços das ações respondem aos fatores que acabam as influenciando. Existem mercados lentos, com baixa volatilidade e mercados rápidos, com alta volatilidade, dependendo muito do mercado que está sendo estudado. Porém, essa rapidez com que o mercado se ajusta aos eventos que o influenciam, pode ser quantificada, o que permite, dessa forma, que qualquer previsão seja bem mais confiável.

Diante do que foi exposto, ao pensar no mercado financeiro, é preciso considerar sua volatilidade, já que o preço das ações são variáveis e existem inúmeros fatores que o influenciam.

2.2 Previsão de Séries Temporais

Uma das características que são apresentadas pelos dados temporais é a relação entre o resultado das observações e o tempo, isso faz com que a análise e a representação desses dados apresentem características específicas que diferem da análise de dados tradicionais, o que traz vantagens que podem contribuir para a descoberta de novos conhecimentos dos dados estudados (AIKES, 2012).

Segundo Reis (2007) uma série temporal pode ser definida como um conjunto de observações regulares ao longo do tempo de uma determinada variável, que pode ser referente a temperaturas, vendas mensais e até mesmo fechamento diário de uma ação na bolsa de valores. Ou seja, é um sistema casual que dependendo da variável a ser estudada pode ter uma constância maior ou menor em relação ao tempo, mas que cria padrões que podem ser identificados e estudados a partir de processos estatísticos e gráficos.

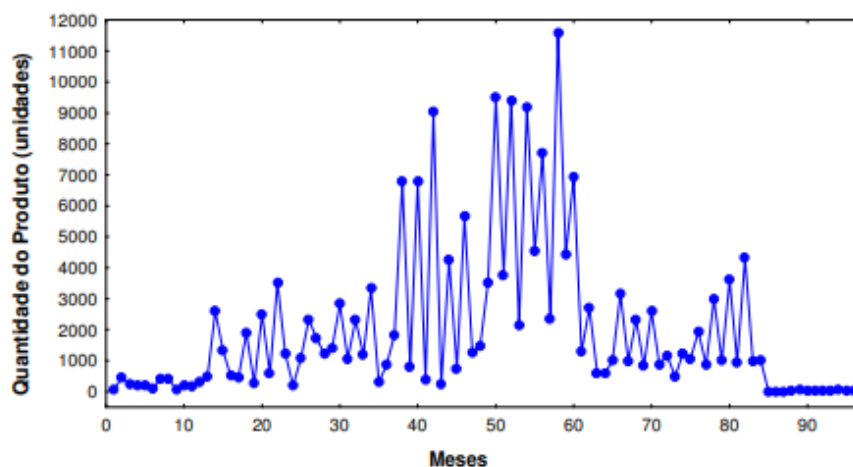
De acordo Reis (2007), a partir do modelo clássico, toda série temporal pode ser baseada em quatro padrões, pois não é preciso necessariamente apresentar todos os componentes, que são:

- Tendência, representa a direção que os dados estão seguindo ao longo do tempo, aumentando, diminuindo ou permanecendo estáveis de forma geral.
- Variações cíclicas ou ciclos, flutuação de valores de longo prazo que não podem ser consideradas sazonalidade devido ao período de tempo maior.
- Sazonalidade, variações regulares e previsíveis que acontecem com intervalos fixo de tempo.
- Variações irregulares ou ruídos, são caracterizados por flutuações não sistemáticas e aleatórias da série temporal, causadas por eventos aleatórios ou fatores imprevisíveis.

Conforme citado por Morettin (1981), depois que uma determinada série temporal foi obtida através dos resultados das observações em seus respectivos tempos, podemos investigar o mecanismo gerador da série temporal, fazer previsões futuras da série, descrever o comportamento da série, procurar periodicidade relevante nos dados, pois independente do que está sendo buscado será preciso a utilização de modelos probabilísticos ou estocásticos.

Figueiredo (2008), descreve que uma série temporal não é algo complexo, mas quando o objetivo é estudá-la, será necessário que esteja diretamente relacionada com cálculos estatísticos, representação gráfica e ajustamento de características simples, como apresentado na Figura 1.

Figura 1 – Representação gráfica de uma série temporal



Fonte: Figueredo (2008, p.34)

A modelagem de uma série temporal consiste em identificar o processo gerador daquela série, sendo que tal procedimento está relacionado com estimação de parâmetros do modelo do processo e avaliação da qualidade do diagnóstico, enquanto a previsão envolve em estimar valores futuros e, dependendo do que está sendo analisado, pode trazer vários benefícios, além de permitir uma visão antecipada da direção que a série provavelmente tomará (FIGUEREDO, 2008).

A utilização de séries temporais para realizar previsões parte do princípio de que dados futuros estejam relacionados com os dados passados, não sendo dependentes de outras variáveis. É considerado um método simples, porém quando feito da forma correta pode gerar resultados muito satisfatórios. Um dos primeiros passos para montar um modelo de previsão é plotar os dados e identificar quais os componentes estão ligados às características da série temporal (TUBINO, 2006).

2.3 Redes Neurais

Em 1943 foi o ano em que aconteceu o primeiro estudo relacionado às Redes Neurais, nele foram analisadas as formas como o neurônio biológico se comporta, com o objetivo de criar um modelo matemático. As conclusões geradas por esse trabalho foram muito relevantes para inovações computacionais (CARDON, 1994).

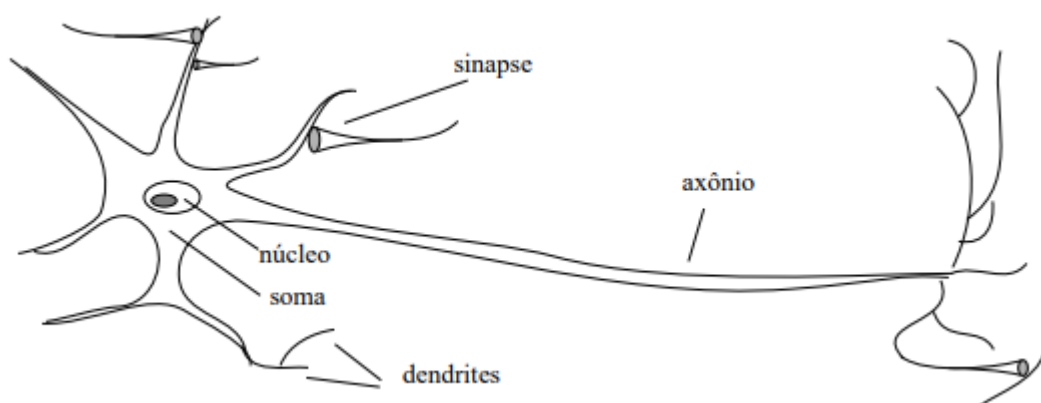
O cérebro humano é um sistema extremamente complexo estudado pelo homem e muitas descobertas sobre ele ainda serão feitas (BAELE *apud* MUELLER, 1996).

Schossler (2017) nos diz que o sistema nervoso é formado por uma rede de neurônios e cada um deles tem a capacidade de receber a informação, processá-la e se comunicar com milhares de outros neurônios paralelamente. A estrutura, representada pela Figura 2, se divide em três partes: o corpo da célula, o axônio e os dendritos, essas partes têm funções complementares, porém cada uma tem sua função específica.

Os dendritos são responsáveis por receber as informações que são os impulsos nervosos e os transportam em direção ao corpo celular. No corpo celular, essas informações são processadas e novos impulsos são gerados. Os novos impulsos são conduzidos pelo axônio até suas terminações e assim podem ser transmitidos para os neurônios da célula alvo, através da sinapse. A partir dela, que se formam os pontos de comunicação entre um neurônio e outro, ou seja, é através das sinapses que os milhares de neurônios se transformam em uma rede neural (SCHOSSLER, 2017).

A rede neural é composta por vários neurônios, geralmente o modo como cada neurônio opera de maneira individual, está relacionada com a combinação linear das entradas e o respectivo valor de cada uma, seguida pela aplicação de uma função de ativação. As limitações ligadas aos tipos de redes e algoritmos de aprendizagem possível estão diretamente ligadas às características do problema que deve ser resolvido. Após a estrutura da RNA ser definida, é necessário treiná-la, além disso, para que seja capaz de solucionar a tarefa proposta de forma eficiente é preciso adequar os graus de liberdade de maneira ótima (Rauber, 2005).

Figura 2 – Neurônio biológico



Fonte: Rauber (2005, p.4)

As redes neurais artificiais (RNA) se baseiam na maneira como o cérebro humano processa e transmite as informações e a partir desse método, aplicá-lo na resolução de problemas (BEALE *apud* MUELLER, 1996).

Como as redes neurais artificiais são fundamentadas nas redes neurais biológicas várias características que são comuns no processo cognitivo humano (WASSERMAN *apud* MUELLER, 1996). Algumas características podem ser destacadas, como: a capacidade de adaptação por meio da experiência, que é melhorar o processo de conhecimento e desempenho, não somente por tentativa e erro, mas também por um processo de treinamento, ou seja, uma rede neural tem a habilidade de aprimorar e aprender com a prática. E, a capacidade de generalizar o conhecimento, não se limitando a situações específicas, podendo aplicar o conhecimento recebido em cenários que antes eram totalmente desconhecidos (ALBANEZ, 2017).

O modo como as redes neurais solucionam problemas se diferencia em alguns aspectos em relação aos métodos estatísticos, o principal deles é o fato de as RNs não fazerem suposições específicas sobre a distribuição ou propriedade dos dados, o que as torna versáteis e aplicáveis em uma grande variedade de situações práticas. Além disso, são fundamentalmente não-lineares em relação ao modo como tratam os problemas, ou seja, são extremamente eficazes na modelagem de dados com padrões complexos. As redes neurais são muito variadas, já que cada uma pode ser projetada com objetivos, estruturas e algoritmos específicos (SOUZA, 2008).

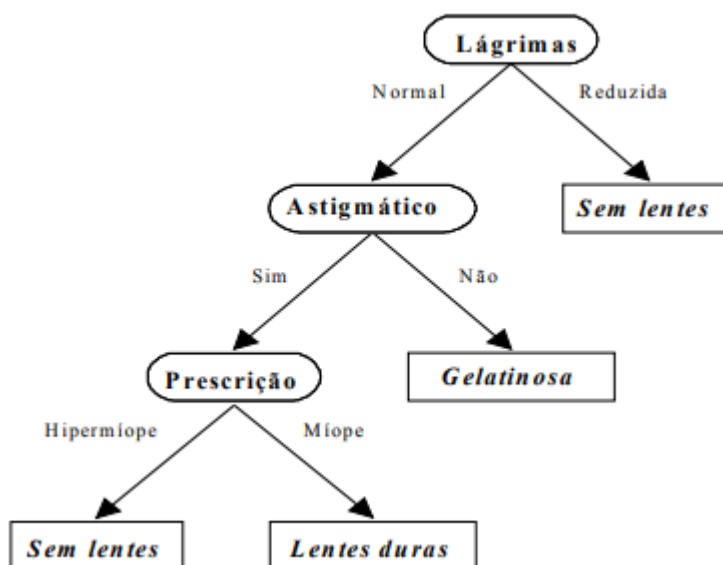
2.4 Árvores de Decisão

A árvore de decisão, também pode ser chamada de árvore de classificação, é uma representação gráfica que mostra as consequências de decisões presentes e futuras. Como ponto de partida, ela recebe o conjunto de informações que descrevem o objeto ou situação, e como saída sua decisão, ou seja, o valor de saída está totalmente relacionado com o conjunto de atributos de entrada. Tanto as características de entrada quanto os valores de saída podem ser de natureza discreta ou contínua. Quando o processo se trata de aprender a partir de dados de valores discretos é uma aprendizagem de classificação, caso seja um processo de aprendizagem a partir de uma função de valores contínuos, é denominada aprendizagem de regressão (NORVIG *apud* MARIN, 2013).

A estrutura de uma árvore de decisão é organizada em níveis que consistem em nós e conexões entre eles. O nó que está localizado no topo da árvore é chamado de nó raiz, que é o ponto de partida. A partir do nó raiz, ramificam-se vários caminhos, que são conhecidos como links ou ramos, que tem como função fazer a ligação com os demais nós. E por fim, as soluções são representadas pelas folhas que também recebem o nome de nós terminais (TAKATUZI *apud* MITCHELL, 1997).

A Figura 3, apresenta uma árvore de decisão relacionada com o uso de lentes de contato.

Figura 3 – Exemplo de árvore de decisão



Fonte: WITTEN, FRANK (2005, p.19)

Neste caso, a árvore inicia com um teste na taxa de produção e lágrimas, e os dois primeiros ramos correspondem a dois resultados possíveis. Se a taxa de produção de lágrimas estiver reduzida, o resultado não utilizar lentes, caso a produção de lágrimas estiver normal, é realizado um segundo teste em relação ao astigmatismo. No final, independentemente dos resultados dos testes anteriores, sempre chegaremos a uma folha de árvore (WITTEN; FRANK, 2005).

Segundo (TAKATUZI *apud* MITCHELL, 1997), o aspecto fundamental desses algoritmos se encontra na necessidade de estabelecer um critério de que selecione o melhor atributo que dividirá o conjunto de exemplos em subconjuntos mais semelhantes a cada passo do algoritmo. Para criar árvores de classificação mais compactas e que proporcione um resultado preciso, é extremamente importante

selecionar algum atributo que gere nós mais “puros”. Um nó somente é considerado “puro” quando todos seus exemplos de treinamento são da mesma classe. Com isso, um nó pode ser considerado “impuro”, caso seus exemplos de treinamento sejam de classes diferentes.

De acordo com (MARIN, 2013) computacionalmente as árvores de decisão são consideradas uma forma muito eficaz e extremamente utilizada para a resolução de problemas de reconhecimento de padrões. Entre as razões que justificam o alto número de utilizações desse método está o fato de que, o conhecimento obtido pode ser descrito através de regras, que podem ser transmitidas por meio de linguagem natural, proporcionando que a interpretação e entendimento dos resultados seja mais simples. Além disso, é um método que pode ser aplicado à base de dados muito grandes, devido a sua indução que é computacionalmente rápida.

2.5 Correlação

De acordo com Virgillito (2017), Karl Person através de estudos matemáticos, realizou contribuições muito importantes para a Estatística. Dentre eles, por volta de 1896, determinou o algoritmo que mede o grau de correlação simples entre duas variáveis.

A fórmula que determina o grau de correlação é representada pela Equação 1 (p.358).

$$r_{x,y} = \frac{\sum(X.Y) - \frac{(\sum X)(\sum Y)}{n}}{\left[\sqrt{\left(\sum X^2 - \frac{(\sum X)^2}{n} \right)} \cdot \left[\sum Y^2 - \frac{(\sum Y)^2}{n} \right] \right]} \quad (1)$$

Em que:

n = número de variáveis observadas;

$\sum(X.Y)$ = somatória dos resultados das multiplicações dos valores de X pela respectiva variável Y ;

$\sum X . \sum Y$ = soma dos valores da variável X multiplicado pela soma dos valores da variável Y ;

$\sum X^2$ = somatória dos quadrados da variável X ;

$(\sum X)^2$ = quadrado da somatória dos valores da variável X ;

$\sum Y^2$ = somatória dos quadrados da variável Y ;

Pearson concluiu que independente da variável estudada, o grau de correlação, conhecido como coeficiente de correlação, sempre estará entre dois valores:

$$-1 \leq r_{x,y} \leq +1$$

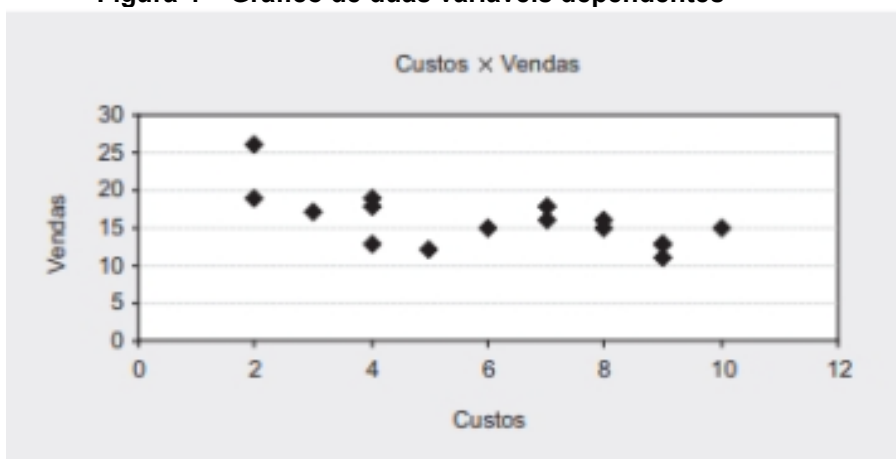
Sendo que, caso o valor de $r_{x,y} = -1$ existe uma correlação perfeitamente negativa. E para quando o valor de $r_{x,y} = +1$ existe uma correlação perfeitamente positiva. Ainda, se $r_{x,y} = 0$ a correlação entre as variáveis é perfeitamente nula ou não existente.

2.6 Regressão

A presença de conexões entre variáveis pode oferecer, em determinadas ocasiões, informações úteis para a tomada de decisões. No entanto, tais decisões estão relacionadas a dados do passado, gerando uma certa incerteza quanto à capacidade desses dados em prever de forma isolada o comportamento futuro das variáveis em estudo. Considerando que as estratégias que foram utilizadas anteriormente sejam aplicadas novamente no futuro, o método de regressão busca modelar o comportamento passado das variáveis em questão e, por meio de uma equação descritiva, projetar como esse comportamento se desdobrará no futuro. (VIRGILLITO, 2017).

Ainda, segundo Virgillito (2017), os matemáticos Gauss e Legendre foram os responsáveis pela formulação dessa técnica por volta de 1806. Posteriormente, Galton, ao alterar seu significado, estabeleceu as bases para que Pearson, ao compreender as relações estabelecidas, pudesse conceber seu coeficiente de correlação. Ou seja, o método busca possibilitar a modelagem do comportamento passado da relação entre as variáveis estudadas, permitindo assim estimar o comportamento futuro com base nesses padrões.

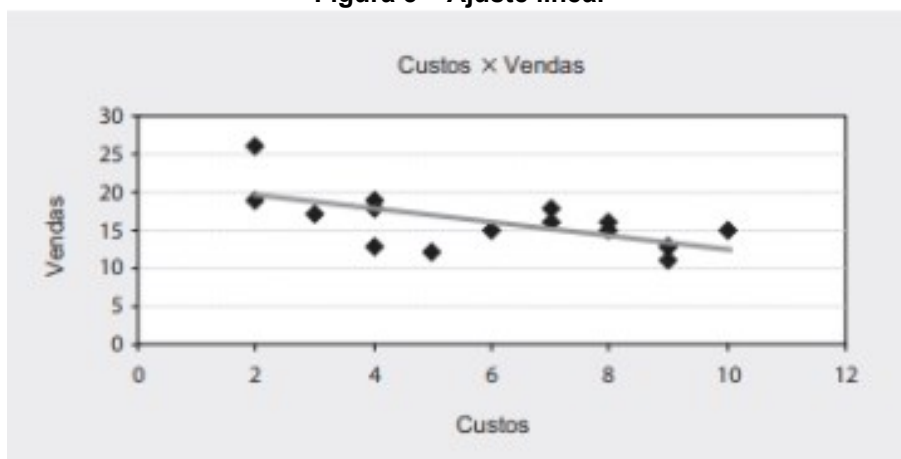
A representação de duas variáveis dependentes é apresentada no gráfico de dispersão representado pela Figura 4.

Figura 4 – Gráfico de duas variáveis dependentes

Fonte: VIRGILLITO (2017, p.368)

A partir do gráfico, é possível determinar a equação que melhor representa o comportamento das variáveis estudadas. Essa equação recebe o nome de curva de ajustamento. As curvas de ajustamento podem ser de diferentes tipos de relação, como: linear, exponencial, logarítmica, entre outras (VIRGILLITO, 2017).

O ajuste linear dos dados do Figura 4 é representado na Figura 5.

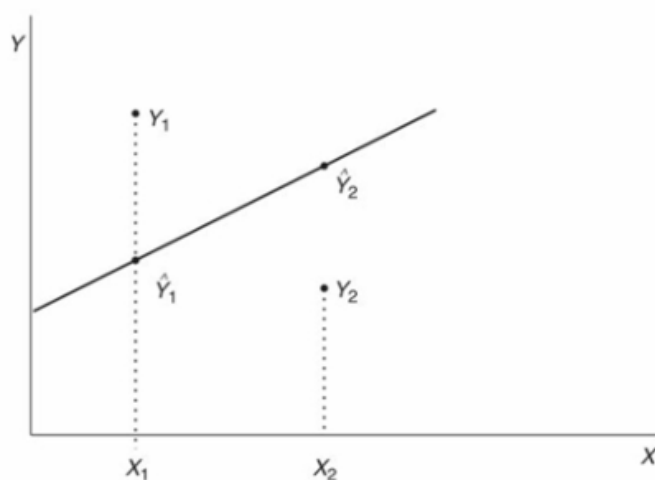
Figura 5 – Ajuste linear

Fonte: VIRGILLITO (2017, p.369)

É possível observar que a curva de tendência indica uma correlação negativa, ou seja, conforme os custos a quantidade de vendas tende a diminuir. Diversos métodos podem derivar uma equação para representar o comportamento de variáveis, no entanto é necessário um método empírico para selecionar a equação que melhor descreva o comportamento (VIRGILLITO, 2017).

Segundo Moreira (2012), o método gráfico gera algumas imprecisões, já que depende de escolhas humanas. Duas pessoas podem chegar a valores diferentes para os mesmos parâmetros. A solução para essa questão é utilizar um método analítico, nesse caso o mais utilizado é o chamado método dos mínimos quadrados, também conhecido como MMQ. O critério utilizado por esse método pode ser observado na Figura 6.

Figura 6 – Base do método dos mínimos quadrados



Fonte: Moreira (2012, p.299)

De acordo com Moreira (2012), no caso de a curva de tendência ser uma reta, como no exemplo da Figura 6, é possível verificar que cada valor da variável causal, tem dois valores de demanda correspondentes, que seriam o valor real, ou seja, que realmente aconteceu e o valor previsto, resultado da equação da reta. O valor previsto é calculado com a substituição do valor de X na equação da reta.

Considerando Y os valores reais e $\hat{Y}$ como sendo os valores previstos. A equação da reta irá gerar os valores previstos, dessa forma pode ser representada pela Equação 2 (MOREIRA, 2012).

$$\hat{Y} = a + bX \quad (2)$$

Atribuindo que X_i seja qualquer valor da variável causal X, que possui um valor de demanda real Y_i e um valor de demanda previsto $\hat{Y}_i$. O erro de previsão é calculado a partir da diferença entre $(Y_i - \hat{Y}_i)$ para $X = X_i$.

A previsão para qualquer valor de X_i pode ser tanto positiva, quanto negativa, de acordo com a posição da reta em relação ao valor real de Y_i . Ou seja, se os erros relacionados a todos os valores de X_i forem somados a tendência do resultado é zero, já que a tendência da reta é aproximar de forma igual todos os valores que se encontram acima e abaixo da reta.

É possível resolver o problema da soma igual a zero ao elevar ao quadrado cada um dos erros de previsão relacionados a cada X . Dessa forma o resultado será sempre positivo e o erro quadrático total pode ser representado pela Equação 3 (MOREIRA, 2012).

$$\sum (Y_i - \hat{Y}_i)^2, \text{ para } i = 1, 2, 3, \dots, n \quad (3)$$

O princípio do método dos mínimos quadrados pode ser definido como encontrar os coeficientes de a e b da reta, que tornem o erro total o menor possível.

Ao derivar a fórmula do erro é possível encontrar as equações conhecidas como Equações normais 4 e 5 (MOREIRA, 2012).

$$\sum Y = na + b\sum X \quad (4)$$

$$\sum XY = a\sum X + b\sum X^2 \quad (5)$$

No caso das equações normais, os índices i foram omitidos. Com exceção dos valores de a e b , todos os outros valores são conhecidos, pois partimos de pares conhecidos (X_i, Y_i) , com n sendo o número desses pares.

3 MÉTODOS E TÉCNICAS

Quanto a natureza da pesquisa, no que diz respeito à forma de abordagem do problema temos uma pesquisa quantitativa, pois será feito um estudo de comparação entre duas variáveis a partir de métodos estatísticos e análise de dados.

De acordo com Prodanov e Freitas (2013, p.70):

Essa forma de abordagem é empregada em vários tipos de pesquisas, inclusive nas descritivas, principalmente quando buscam a relação causa-efeito entre os fenômenos e também pela facilidade de poder descrever a complexidade de determinada hipótese ou de um problema, analisar a interação de certas variáveis, compreender e classificar processos dinâmicos experimentados por grupos sociais, apresentar contribuições no processo de mudança, criação ou formação de opiniões de determinado grupo e permitir, em maior grau de profundidade, a interpretação das particularidades dos comportamentos ou das atitudes dos indivíduos.

Quanto aos objetivos, a pesquisa é de natureza descritiva, pois será realizada a análise dos dados, o registro do método utilizado e a interpretação dos resultados. Triviños (1997) *apud* Gerhardt e Silveira (2009), cita que “A pesquisa descritiva exige do investigador uma série de informações sobre o que deseja pesquisar. Esse tipo de estudo pretende descrever os fatos e fenômenos de determinada realidade”.

Quanto ao método da pesquisa adotada será *ex-post-facto*. Como cita, Prodanov e Freitas (2013, p.65):

Nessa pesquisa, buscamos saber quais os possíveis relacionamentos entre as variáveis. Ela apresenta uma análise correlacional e é aquela que acontece após o fato ter sido consumado, mostra a falta de controle do investigador sobre a variável independente, fato que a diferencia da experimental, sendo, também, muito adotada nas ciências da saúde.

Quanto às etapas metodológicas, essa pesquisa se dividiu em 4 etapas. Na primeira etapa será feita uma fundamentação teórica sobre bolsa de valores, mercado financeiro e preço de ações. Na segunda etapa, será realizado um levantamento de dados sobre o histórico do valor da ação em cada bolsa de valores. Na terceira etapa, será feita a aplicação do método de análise de dados nos valores obtidos. Por fim, na quarta etapa, examina-se os resultados obtidos.

De acordo com Yin (2009), os estudos de caso quantitativos utilizam abordagem estatística e dados numéricos para analisar um fenômeno específico dentro de um contexto detalhado, sendo uma ferramenta poderosa para testar hipóteses e analisar relações causais. Ao combinar métodos quantitativos com

abordagens qualitativas, o estudo de caso se torna uma ferramenta robusta e versátil, capaz de realizar uma pesquisa rigorosa e significativa.

4 RESULTADOS E DISCUSSÕES

As variáveis estudadas serão as ações AMZO34 e AMZN, o dólar e o índice Ibovespa. Os valores considerados para realizar as análises são referentes aos meses de janeiro a junho de 2024, desconsiderando os finais de semana e feriados, tanto do Brasil quanto dos Estados Unidos, gerando um total de 120 dados de cada uma das variáveis.

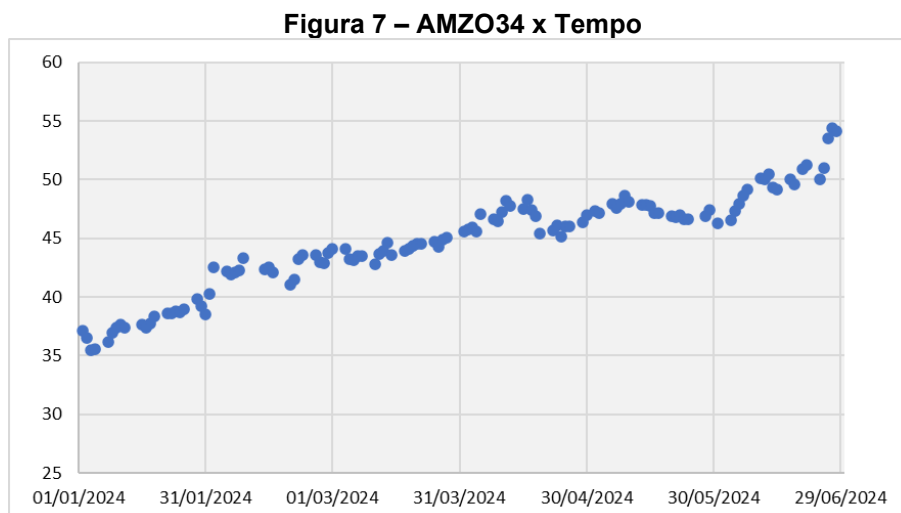
Essa análise é de extrema importância, pois é o ponto de partida que permite compreender o comportamento dos dados antes de realizar uma análise mais aprofundada das variáveis quantitativas utilizadas.

Tabela 1 – Estatísticas descritivas das variáveis quantitativas

<i>Estatísticas</i>	<i>AMZO34</i>	<i>AMZN</i>	<i>DOLAR</i>	<i>IBOVESPA</i>
Média	44,69	175,47	5,08	126864,51
Erro padrão	0,38	1,08	0,02	282,08
Mediana	45,29	178,32	5,03	127596,50
Moda	47,95	178,15	4,87	-
Desvio padrão	4,16	11,85	0,17	3090,08
Variância da amostra	17,28	140,36	0,03	9548597,14
Curtose	-0,28	0,07	0,29	0,06
Assimetria	-0,32	-0,94	1,03	-0,69
Intervalo	18,93	53,28	0,74	13696,00
Mínimo	35,46	144,57	4,85	119138,00
Máximo	54,39	197,85	5,59	132834,00
Soma	5362,31	21056,42	610,13	15223741,00
Contagem	120,00	120,00	120,00	120,00
Nível de confiança(95,0%)	0,75	2,14	0,03	558,56

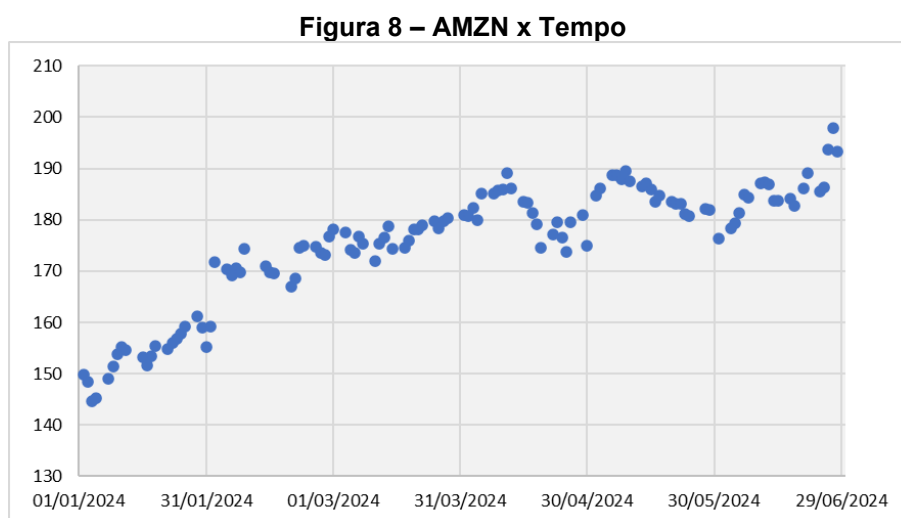
Fonte: Autoria Própria (2025)

A partir das informações apresentadas na Tabela 1, pode-se notar que o dólar foi a variável com menor variação em relação à média, apresentando um desvio padrão de 0,17, enquanto o Ibovespa, que é o índice utilizado para acompanhar o desempenho da bolsa de valores brasileira, foi o que teve maior variação em relação à média, com um desvio padrão de 3090,08, que é justificado pelo fato de os valores máximos e mínimos estarem mais distantes da média, já as variáveis AMZO34 e AMZN possuem um desvio padrão de 4,16 e 11,85, respectivamente.



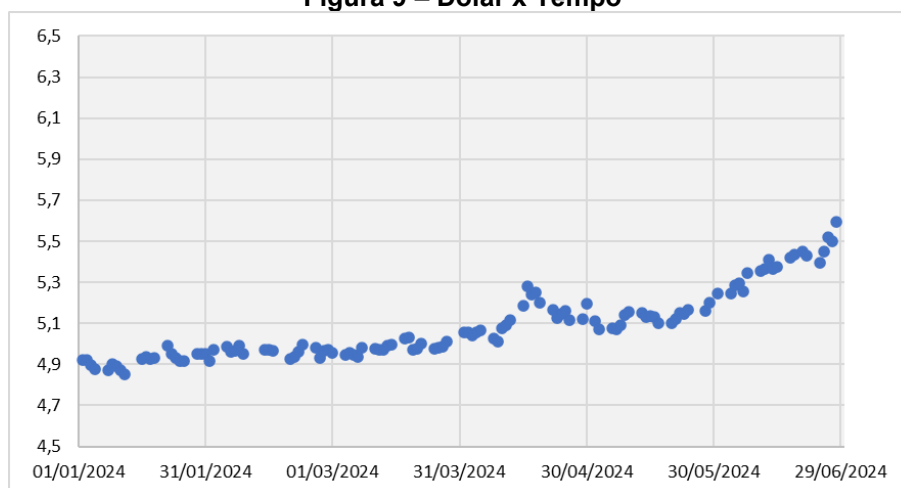
Fonte: Autoria Própria (2025)

A Figura 7 mostra uma tendência crescente ao longo do tempo, com pequenas oscilações que indicam variações pontuais.



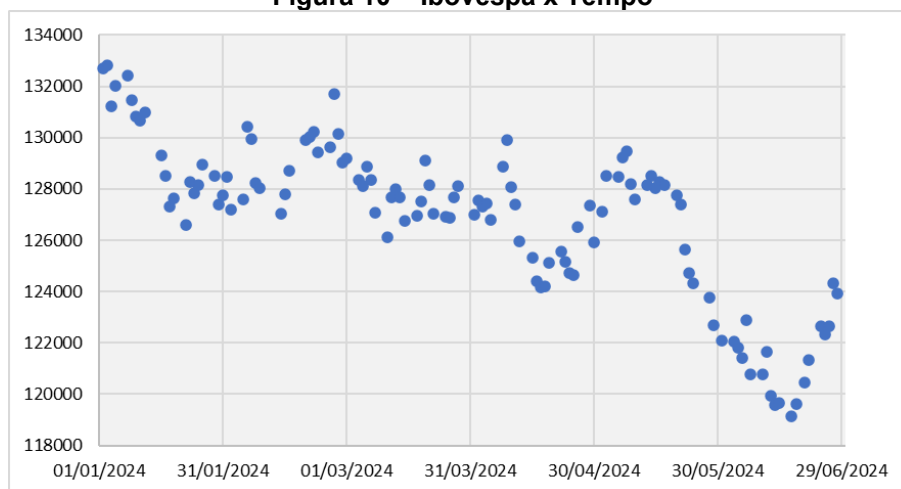
Fonte: Autoria Própria (2025)

A Figura 8 mostra um crescimento geral, com algumas oscilações intermediárias, ou seja, apresenta períodos de estabilidade, algumas quedas e recuperações ao longo do tempo.

Figura 9 – Dólar x Tempo

Fonte: Autoria Própria (2025)

No gráfico da Figura 9 nota-se estabilidade inicial, seguida de um pequeno aumento, e posterior crescimento moderado em relação ao período inicial.

Figura 10 – Ibovespa x Tempo

Fonte: Autoria Própria (2025)

O gráfico da Figura 10 mostra uma certa instabilidade em comparação aos gráficos anteriores, com oscilações mais rigorosas e tendência de queda, porém com sinais de recuperação no último mês.

Essa análise gráfica permite avaliar que os valores das variáveis AMZO34, AMZN e o Dólar, apesar das flutuações ao longo do período estudado, permanecem relativamente estáveis, com tendência de alta e baixa volatilidade, o que pode ser justificado pelo fato do valor da média e da mediana serem muito próximos. Enquanto o gráfico que representa a série temporal do Ibovespa apresenta menor estabilidade

e maior volatilidade em relação as outras variáveis estudadas, além de possuir uma tendência de queda ao decorrer do tempo.

Considerando o estudo de Pearson, foi calculado o grau de correlação das variáveis selecionadas com a variável estudada AMZO34, utilizando o software Excel:

Tabela 2 – Coeficiente de Pearson

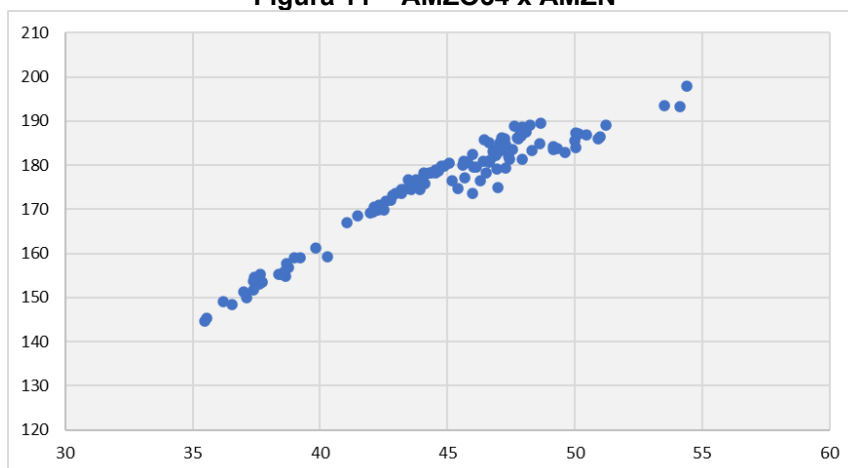
Variável	Grau de Correlação (AMZO34)
AMZN	0,959630655
DOLAR	0,863724423
IBOVESPA	-0,708123815

Fonte: Autoria Própria (2025)

Como é possível observar, as variáveis AMZN e o dólar apresentam um grau de correlação de aproximadamente 0,96 e 0,86, respectivamente, valores muito próximos de 1, ou seja, existe uma relação linear forte e positiva. Enquanto, o Índice Ibovespa apresenta grau de correlação de aproximadamente -0,71, que pelo fato de não estar muito próximo de -1, representa uma relação linear moderada e negativa. Desta forma, as três variáveis foram levadas em consideração para realizar os cálculos da Regressão Linear.

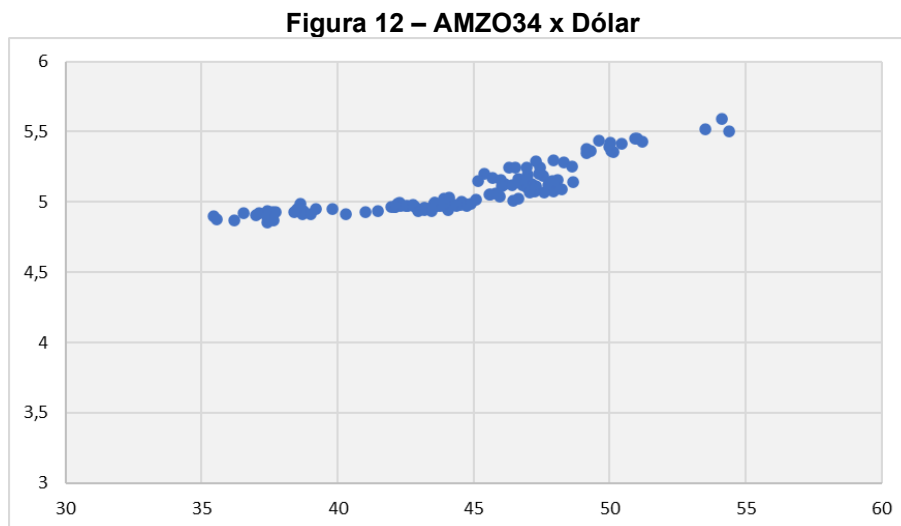
A suposição da linearidade na regressão linear refere-se à relação entre as variáveis independentes (AMZN, Dólar, Ibovespa) e a variável dependente (AMZO34). A suposição implica que a relação entre as variáveis pode ser expressa por uma equação linear, ou seja, o aumento ou redução de uma variável independente leva a um aumento ou redução de uma variável dependente.

Figura 11 – AMZO34 x AMZN



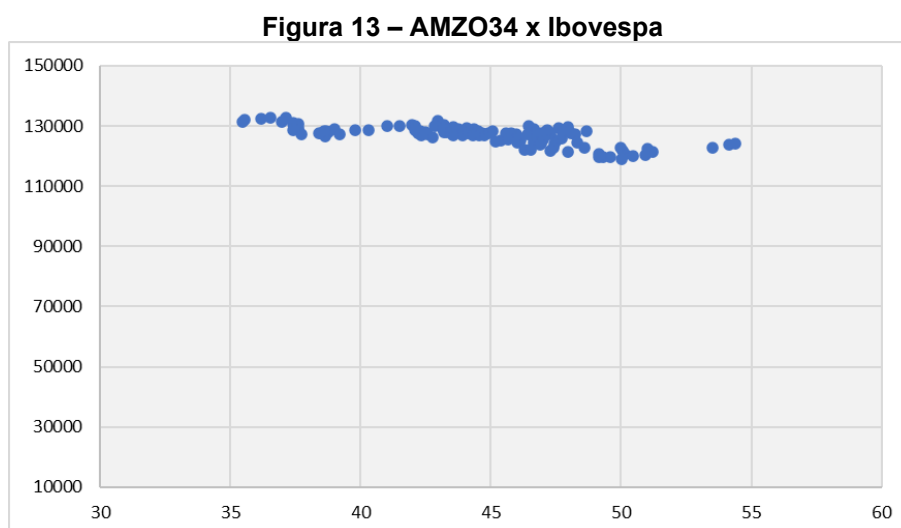
Fonte: Autoria Própria (2025)

A partir do gráfico da Figura 11, a tendência dos dados permite observar o comportamento linear positivo entre as variáveis AMZO34 e AMZN e, não apresenta sinais visíveis de curvatura ou padrões não lineares, o que evidencia que a suposição de linearidade está sendo atendida.



Fonte: Autoria Própria (2025)

Na Figura 12, também é possível observar o comportamento linear positivo entre as variáveis AMZO34 e o Dólar.



Fonte: Autoria Própria (2025)

Na Figura 13, observa-se um comportamento linear levemente negativo entre as variáveis AMAZO34 e Ibovespa.

Diante de todos os dados apresentados, verifica-se que o comportamento de todas as variáveis independentes em relação a variável dependente é linear, obedecendo a suposição de linearidade, que caso fosse violada, poderia gerar estimativas incorretas, reduzir a precisão das previsões e invalidar testes de hipóteses iniciais.

Utilizando o Software Excel, foi aplicado o método de Regressão Linear para descrever adequadamente a relação entre as variáveis e verificar a possibilidade ou não da previsão da variável AMZO34. Os preditores empregados no cálculo foram AMZN, DOLAR e IBOVESPA.

Tabela 3 – Estatísticas de Regressão	
<i>Estatística de regressão</i>	
R múltiplo	0,998382864
R-Quadrado	0,996768343
R-quadrado ajustado	0,996684766
Erro padrão	0,239375637
Observações	120

Fonte: Autoria Própria (2025)

De acordo com a Tabela 3, o resultado obtido a partir da regressão foi um valor de R múltiplo, que mede o grau de correlação entre as variáveis independentes e variável dependente, de aproximadamente 0,998, ou seja, como o valor é muito próximo de 1 indica uma relação muito forte e positiva. O valor de R-quadrado de 0,9968, significa que 99,68% das variações dos valores da ação AMZO34, são explicadas pelas variáveis independentes (AMZN, DOLAR, IBOVESPA). O valor de R-quadrado ajustado muito próximo de 1 indica que o modelo tem uma boa qualidade e as variáveis independentes realmente contribuem para explicar a variável dependente.

Tabela 4 – ANOVA					
	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	3	2050,155819	683,3852728	11926,29982	3,0404E-144
Resíduo	116	6,646880661	0,057300695		
Total	119	2056,802699			

Fonte: Autoria Própria (2025)

A análise da Tabela 4, confirma a boa qualidade do ajuste do modelo, pelo fato do SQ Regressão, que representa a parte da variação que é explicada pelo modelo, ter o valor significativamente maior que o SQ Resíduo, que representa a parte

da variação que não é explicada pelo modelo. Além disso, o fato de o valor de F ser elevado e o valor de F de significação ser extremamente pequeno, indicam que o modelo é altamente significativo.

Tabela 5 – Coeficientes

	<i>Coeficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Interseção	-52,38067167	3,0882392	-16,9613391	3,38E-33
AMZN	0,24309002	0,002595	93,67637096	3,7E-111
DOLAR	9,989095267	0,3066451	32,57542855	3,32E-60
IBOVESPA	2,85571E-05	1,433E-05	1,993469202	0,048557

Fonte: Autoria Própria (2025)

A partir da Tabela 5, observa-se que todos os valores-P são menores que 0,05 e os valores de Stat t são diferentes de 0, indicando que todas as variáveis independentes possuem relevância estatisticamente significativa, ou seja, devem ser consideradas no modelo. Com isso, podemos desenvolver a equação para prever os valores da variável dependente (AMZO34) a partir das variáveis independentes:

$$y = -52,38 + 0,24x_1 + 9,99x_2 + 0,0000285x_3$$

Sendo que:

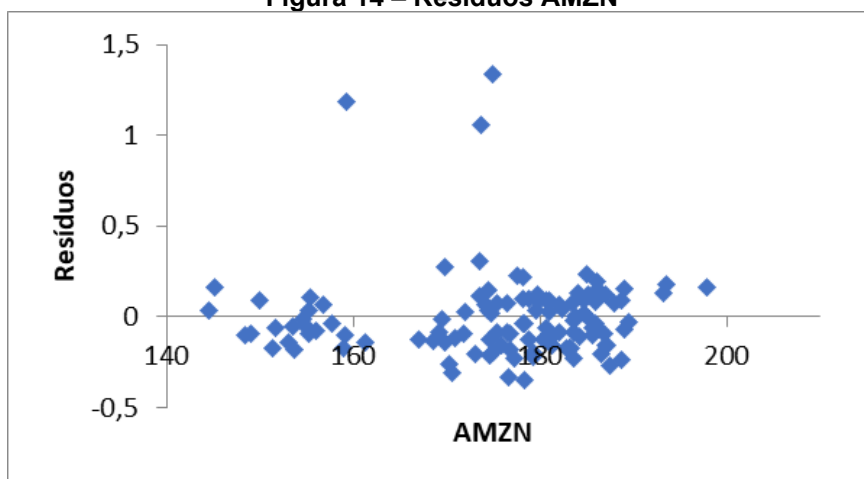
y = valor previsto de AMZO34

x_1 = valor de AMZN

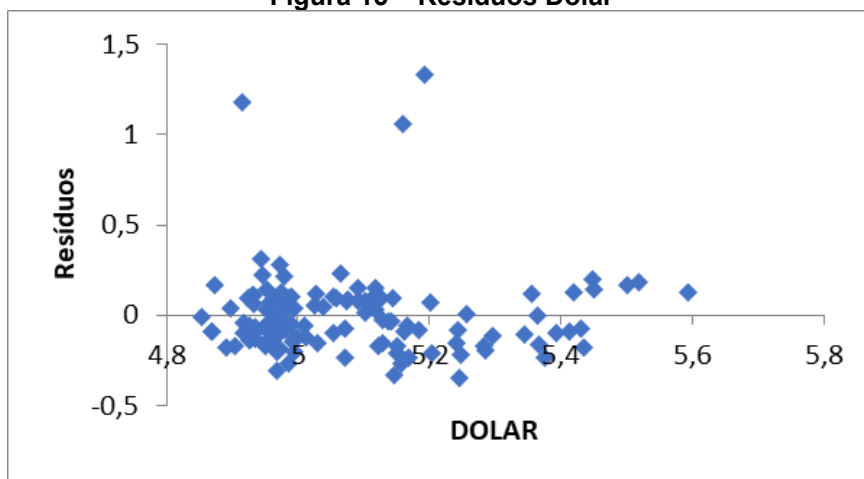
x_2 = valor do Dólar

x_3 = valor do Índice Ibovespa

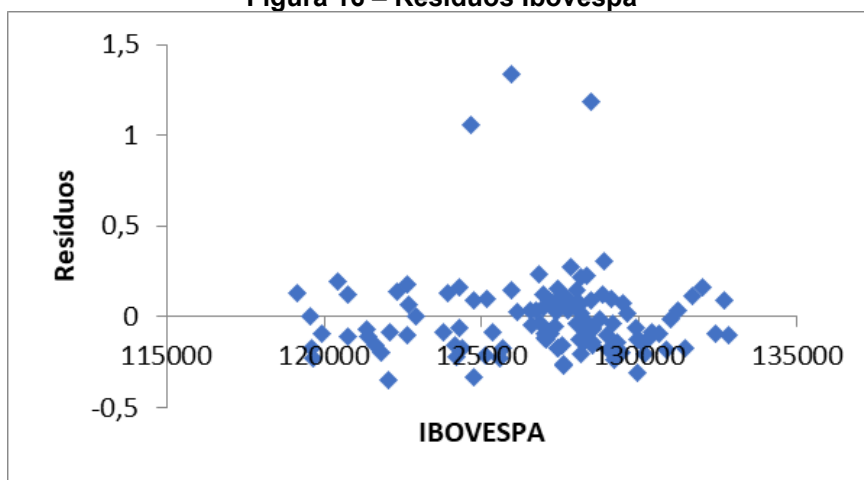
A seguir são apresentados os gráficos de plotagem de resíduos das variáveis independentes resultantes do processo de regressão linear múltipla.

Figura 14 – Resíduos AMZN

Fonte: Autoria Própria (2025)

Figura 15 – Resíduos Dólar

Fonte: Autoria Própria (2025)

Figura 16 – Resíduos Ibovespa

Fonte: Autoria Própria (2025)

Através da análise das Figuras 14, 15 e 16, é possível identificar que os gráficos de resíduos apresentam *outliers*, dados com um desvio significativo do padrão geral, ou seja, valores que estão distantes dos outros pontos em relação às variáveis independentes ou à variável dependente. Esses pontos em um modelo de regressão, podem influenciar de forma desproporcional o resultado da análise, gerando estimativas de coeficientes, interceptos e medidas de ajuste que podem não representar adequadamente o comportamento geral dos dados.

A presença de *outliers* nos dados, citada anteriormente, podem ocorrer por diversos motivos como: erros de medição, variabilidade natural, erros no processo de amostragem, eventos raros, efeitos de intervenção, erro de modelagem, dados falsificados ou manipulados, processamento de dados, problemas de tempo ou sazonalidade.

As possíveis causas da ocorrência de *outliers* na amostra, estão relacionadas com eventos atípicos que aconteceram no mercado financeiro e com a empresa estudada, conforme as seguintes datas:

- 01/02/2024 – a empresa Amazon reportou os resultados referentes ao quarto trimestre de 2023, superando as expectativas do mercado.
- 25/04/2024 – ocorreu uma desvalorização da ação devido a um movimento mais cauteloso do mercado financeiro, que aguardava a divulgação dos resultados trimestrais, que de acordo com as previsões de receitas poderiam ficar abaixo das expectativas do mercado, além de fatores macroeconômicos, como a incerteza do Federal Reserve em relação a taxa de juros.
- 30/04/2024 – o valor das ações apresentou uma nova queda, pelo fato de a divulgação dos resultados trimestrais serem positivos, porém ficaram abaixo do esperado pelos investidores.

Neste caso, a presença de *outliers* no conjunto de dados está relacionado com determinados eventos, mas que fazem parte do fenômeno, ou seja, refletem variações legítimas, portanto não se deve desconsiderá-los.

Porém, a sua ocorrência pode interferir de maneira negativa na análise dos dados. Dependendo do impacto causado no modelo e no contexto da análise é possível realizar o tratamento dos *outliers* para minimizar os danos causados, no entanto, mesmo com a presença de dados com desvio significativo do padrão geral,

o resultado apresentado pela regressão linear múltipla foi satisfatório, conforme os valores apresentados, tornando-se dispensável realizar processos de tratamento.

5 CONSIDERAÇÕES FINAIS

A estatística descritiva é parte essencial da análise de dados, pois a partir dela é possível resumir, organizar e interpretar informações. É extremamente importante para várias áreas do conhecimento, inclusive sendo o passo inicial e fundamental para análises mais complexas, como a regressão linear múltipla, que também é uma ferramenta estatística de grande importância, pela sua capacidade de explicar a relação entre variáveis, realizar previsões e auxiliar na tomada de decisão baseada em dados, tornando uma ferramenta indispensável para a análise de dados.

Nesse trabalho foi realizada a análise da ação AMZO34, buscando verificar sua correlação com as variáveis AMZN, Dólar e Ibovespa, através de um modelo de regressão linear múltipla. Para a análise, foi considerado o histórico de seis meses das variáveis (janeiro de 2024 a junho de 2024), desconsiderando os finais de semana e feriados, tanto do Brasil quanto dos Estados Unidos, pois a bolsa de valores de ambos os países não opera. Tais variáveis passaram por uma nova análise estatística, com objetivo de verificar o nível de correlação entre elas, e por fim identificar quais delas deveriam ser consideradas no modelo de regressão. Constatou-se que todas deveriam ser utilizadas, pois tiveram um valor do coeficiente de Pearson significativo, as variáveis AMZN e Dólar com uma elevada correlação positiva, enquanto a variável Ibovespa, com uma correlação moderada negativa.

A aplicação da regressão linear múltipla, gerou valores de R-quadrado e R múltiplo muito próximos de 1, indicando a boa qualidade do modelo para explicar a variação dos dados e a força de relação entre as variáveis dependente e independentes. Além disso, o fato do valor de F ser elevado, de F significação muito próximo de 0 e os valores-p para todas as variáveis serem menores que 0,05, indicam um modelo altamente significativo, reforçando os resultados de R-quadrado e R múltiplo.

A análise dos gráficos dos resíduos das variáveis, identificou a presença de alguns *outliers*, que em determinadas situações podem impactar negativamente os resultados obtidos pelo modelo. Porém, foi constatado que se tratava de determinados eventos naturais, sendo variações legítimas e que não interfeririam nos valores, dispensando técnicas de tratamento.

Por fim, os resultados obtidos foram satisfatórios, pelo fato de o modelo de regressão apresentar valores estatísticos que demonstrem sua eficiência e qualidade, comprovando que as variáveis possuem correlação entre si.

REFERÊNCIAS

AIKES Junior, Jorge. **Estudo da influência de diversas medidas de similaridade na previsão de séries temporais utilizando o algoritmo KNN-TSP**. Dissertação (Mestrado) - Unioeste, 2012.

ALBANEZ, Daniela de Oliveira. **Redes neurais artificiais aplicadas à segmentação de imagens**. 2017.

CARDON, André; MÜLLER, Daniel Nehme; NAVAUX, Philippe. **Introdução às redes neurais artificiais**. Porto Alegre, 1994.

CHAMBI CONDORI, Pedro Pablo. **La Volatilidade de los mercados financieros globalizados: impacto em la bolsa de valores de Lima - Peru**. Quipukamayoc, [S.I.], v. 25, n. 47, p. 103–111, 2017. Disponível em: <<https://revistasinvestigacion.unmsm.edu.pe/index.php/quipu/article/view/13808>>. Acesso em: 10 de Nov de 2023.

FIGUEREDO, Clodoaldo José. **Previsão de Séries Temporais Utilizando a Metodologia box and Jenkins e Redes Neurais para Inicialização de Planejamento e Controle de Produção**. Dissertação (Mestrado) - UFPR, 2008.

GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo. **Métodos de pesquisa**. Plageder, 2009.

GUEDES, Rodrigo Boaventura. **A ascensão da participação jovem na bolsa de valores: fatores de risco e perspectivas**. 2021. 38 f. Trabalho de Conclusão de Curso (Bacharelado em Administração) - Universidade de Brasília, Brasília, 2021.

MARIN, Maikon Aloán. **Indução de Árvores de Decisão para a Inferência de Redes Gênicas**. 2013. Tese de Doutorado. Tese (Doutorado) - Tese de Doutoramento, Universidade Tecnológica Federal do Paraná.

MEDEIROS, Marcos Vinicius et al. **Ausência do brasileiro no mercado financeiro**. South American Development Society Journal, [S.I.], v. 6, n. 17, p. 211,

ago. 2020. ISSN 2446-5763. Disponível em:
<<http://www.sadsj.org/index.php/revista/article/view/329>>. Acesso em: 08 dez. 2022.

MITCHELL, Tom M. **Machine Learning**. McGraw-Hill, New York, 1997.

MOREIRA, Daniel A. **Administração da Produção e Operações – 2ª edição Revista e Ampliada**. 2. Ed. São Paulo: Cengage Learning Brasil, 2012. *E-book*. p.300. ISBN 9788522110193. Disponível em:
<<https://integrada.minhabiblioteca.com.br/reader/books/9788522110193/>>. Acesso em: 15 nov. 2023.

MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Modelos para previsão de séries temporais**. Rio de Janeiro: Instituto de matemática pura e aplicada, 1981.

MUELLER, Alessandro. **Uma aplicação de redes neurais artificiais na previsão do mercado acionário**. 1996. Diss. Dissertação (Mestrado em Engenharia de Produção) - UFSC, Santa Catarina, 1996.

PAIVA, R. T.; SILVA, H. A.; SOUZA, J. C. M. de.; NOVÔA, N. F.; PEREIRA, C. M. M. de A. **O perfil do investidor individual no mercado financeiro**. Revista Vianna Sapiens, [S. l.], v. 11, n. 2, p. 30, 2020. Disponível em:
<<https://www.viannasapiens.com.br/revista/article/view/694>>. Acesso em: 8 dez. 2022.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. 2. Ed. Novo Hamburgo: Feevale, 2013.

RAUBER, Thomas Walter. **Redes neurais artificiais**. Universidade Federal do Espírito Santo, v. 29, 2005.

REIS, M. M. **Análise de séries temporais**. [S.l.], 2007.

SCHOSSLER, Jackson Luis. **Previsão de valor de um ativo financeiro utilizando redes neurais artificiais para processamento de dados de séries temporais**.

2017. 57 f. Trabalho de conclusão de curso (Graduação) - Universidade Tecnológica Federal do Paraná, Medianeira, 2017.

SERASA. Disponível em: <<https://www.serasa.com.br/blog/como-funcionam-os-precos-das-acoes-no-mercado-financieiro/>>. Acesso em: dez.2023.

SOUZA, Rodrigo Clemente Thom de. Previsão de séries temporais utilizando Rede Neural Treinada por Filtro de Kalman e Evolução Diferencial. 2008. 82 f. Dissertação (Mestrado) - Pontifícia Universidade Católica do Paraná, Curitiba, 2008.

TUBINO, D. F. **Manual de planejamento e controle de produção**. 2006.

VIRGILLITO, Salvatore B. **Estatística Aplicada**. Editora Saraiva, 2017. E-book. ISBN 9788547214753. Disponível em: <<https://integrada.minhabiblioteca.com.br/#/books/9788547214753/>>. Acesso em: 15 nov. 2023.

WITTEN, Ian H; FRANK, Eibe. **Data Mining: pratical machine learning tools and techniques**. Morgan Kaufmann Publishers. Second Edition, 2005.

YIN, Robert K. **Case Study Research: Design and Methods**. 4. ed. Thousand Oaks: SAGE Publications, 2009.